Interpreting the Raw Score Tab

Provided by:
Test Validation & Construction Unit
California State Personnel Board

The raw score tab printout provides information to assess the quality of an exam and to establish pass points. It is important to request the appropriate raw score for your needs. If you will be setting one overall pass point for an exam, order the raw score tab for the exam as a whole (this is the usual raw score tab printout). This printout will contain candidate scores for the entire exam. If you will be setting pass points for individual exam sections, order the raw score tab by exam segments. This printout can be requested on the screen that is currently used for ordering the usual raw score tab (screen RR). This printout will contain candidate scores for each of the exam segments. Regardless of the raw score tab that is ordered, the printout is divided into the following four sections: frequency distribution, exam statistics, adverse impact statistics, and frequency histogram. Descriptions of these four sections follow.

Frequency Distribution

The frequency distribution contains specific data regarding the number of candidates that received each raw score in the distribution of scores. Depending on the printout that was ordered, this section will either contain scores for the entire exam or scores for the exam segment.

By reviewing the frequency distribution, the difficulty of the exam or exam segment can be assessed. The frequency distribution can also be used to locate any "breaks" in the distribution. Breaks, or areas in the distribution where candidates did not achieve scores, may be a consideration when setting a pass point. For example, if there is a "break" at, or very close to, the MAC (minimal acceptable competence) level, that may be a determining factor in selecting the pass point.

The frequency distribution consists of five columns.

Maximum Score

The first column is labeled with the maximum score possible on the exam or exam segment. This column contains every score that was achieved by the candidates.

% R

Each value in this column is the percentage correct score that corresponds to the raw score values to the immediate left.

Freq

This column contains the number of candidates that achieved each raw score.

Cum F

This column indicates the total number of candidates that attained the particular raw score or a higher score on the exam. In other words, this value represents, at any given score, the total number of candidates that have been accounted for. When used in setting a pass point, it provides

the total number of candidates which would pass the exam if the pass point were set at that particular score.

C F %
Each number in this column indicates the percent of candidates that achieved that raw score or a higher score on the exam. In other words, this value represents, at any given score, the percentage of candidates that have been accounted for. When used in setting a pass point, it represents the percentage of candidates which would pass the exam if the pass point were set at that particular score.

Exam Statistics

The raw score tab for the exam as a whole will contain statistics for the entire exam and statistics for any exam segments. The raw score tab for exam segments will only contain statistics for the exam segments and not for the entire exam. If the exam segments contain more homogeneous sets of items than the exam as a whole, or if pass points will be set for exam segments, it is more appropriate to interpret the statistics for the exam segments than for the entire exam. The following calculations are included in this section: actual 70%, median, mean, variance, standard deviation, reliability, standard error, choices per item, and chance score.

Actual 70%
This number is the raw score that corresponds with getting 70% correct on the exam or exam segment. Seventy percent should not automatically be used as a pass point. It bears no relationship to any of the job-related, technical, or administrative considerations that should be taken into account. It is important to remember that pass points set above actual 70% must be approved by the Test Validation and Construction Unit of the State Personnel Board.

Median
The median indicates the middle score in the distribution. The median is one of two measures of central tendency that is on the raw score tab printout. Measures of central tendency are indicators of a distribution's average or typical score. If the distribution contains outliers (i.e., extreme scores), the median is the measure of central tendency that should be considered.

Mean
The mean indicates the average score in the distribution. The mean is the second measure of central tendency that is on the raw score tab printout. Measures of central tendency are indicators of a distribution's average or typical score. If the distribution is free of outliers (i.e., extreme scores), the mean is the measure of central tendency that should be considered.

Variance
The variance indicates the amount of dispersion or spread in a distribution. In other words, distributions with greater variance have candidates spread out

between *many* scores instead of having candidates spread out between *fewer* scores. In general, examinations with greater variance are preferred over those with less variance. Examinations with greater variance allow finer distinctions to be made between candidates.

Standard Deviation

The standard deviation is a measure of the dispersion of the exam scores about the mean score. In effect, the standard deviation is the average of the difference of the candidates' scores from the mean score. The larger the standard deviation, the more the scores differ from each other. The standard deviation will be 0.00 if all of the scores are the same. As a rule of thumb, a good standard deviation (that is, one that gives a good distribution of scores) is at least 10% of the length of the exam or exam segment.

Reliability

The reliability of the exam or exam segment is calculated using Cronbach's Coefficient Alpha. The reliability of an exam is the extent to which scores earned on that exam are precise or stable indicators of the candidates' true level of knowledge or skill. When exam reliability is low, there is an increased chance of accepting candidates who cannot perform the job or eliminating candidates who can perform the job. If an exam's reliability coefficient is low, adjustments to the pass point may be necessary to allow for greater error variance.

Reliability coefficients in the range of .90 or above are considered very good; coefficients in the range of .80 to .89 are considered good; and, coefficients in the range of .70 to .79 are only fair. Reliability coefficients are affected by the number of items in the exam or exam segment, therefore lower reliabilities are often found for exam segments which contain fewer items. Consider setting a more lenient pass point if your exam or exam segment has marginal reliability.

Standard Error

The standard error provides an estimate of the accuracy of a candidate's score on the exam or exam segment. The standard error is based on the reliability coefficient. As the reliability of an exam section increases, the standard error decreases. The smaller the standard error of measure, the more confidence you can place in the candidate's score. The standard error is useful when determining how far to stray from the MAC level when setting pass points. In general, pass points should be set no lower than two standard error units below the MAC level.

Choices Per Item

The choices per item indicates the number of alternatives a candidate chooses from on a multiple choice exam. It is important to verify that the number of alternatives is correct for each exam segment.

Chance Score

The chance score indicates the score a candidate would receive on the exam or exam segment by merely guessing. Pass points should always be set well above the chance score.

Adverse Impact Statistics

Consideration should be given to minimizing adverse impact when setting the pass point. A fundamental principle underlying the *Uniform Guidelines* is that employer policies or practices which have an adverse impact on the employment opportunities of any race, sex, or ethnic group are illegal unless justified by business necessity. It is important to remember that adverse impact is tolerated when an exam is valid, the pass point is job-related, and there are no other equally valid exams that would yield less adverse impact. A thorough job analysis is one way of establishing validity.

The *Uniform Guidelines* use a "rule of thumb" as a practical means of determining adverse impact. This rule, known as the "4/5" or "80%" rule, states that a selection rate (passing rate) for any race, sex, or ethnic group which is less than 4/5 or 80% of the rate for the group with the highest selection rate is evidence of adverse impact.

The raw score tab calculates adverse impact statistics for ethnicity, gender, and disability. When setting pass points by exam segments, the adverse impact for each segment should be interpreted rather than the adverse impact of the exam as a whole. It is important to note that adverse impact is calculated for every score. The fact that the exam has adverse impact at one particular score does not indicate that the exam will also have adverse impact at higher scores.

The Rule of One is often used to determine the practical significance of any adverse impact rates detected using the 80% rule. If adverse impact is negated by having one more candidate pass the exam, the adverse impact has no practical significance and can be ignored. The raw score tab does not calculate the Rule of One and therefore you must perform this calculation for any pass point you are considering.

Ethnicity

Adverse impact statistics are calculated for the following ethnic groups: White, Black, Hispanic, Asian, Filipino, Pacific Islander, and Native American. The adverse impact table is divided into the following sections: percent, raw score, % R, freq, cum freq, cum frq %, and counts by ethnic code.

Percent The percent section is comprised of two rows and nine columns of data for each raw score value. The columns represent the various ethnic groups along with the "other" and "DNS" (did not specify) categories. The following illustrates the percent section of the printout.

----------PERCENT----------								
WHITE	**BLACK**	**HISP**	**ASIAN**	**FLIP**	**PACI**	**NATAM**	**OTHER**	**DNS**
68.8	68.8	81.3	71.5	69.5	90.0	0.0	50.0	76.1
76.4	76.4	90.3	79.4	77.2	100.0	0.0		

The first row of data illustrates the percentage of candidates within each ethnic group that received that raw score or a higher raw score on the exam. Also illustrated, is the percentage of candidates that selected "other" as their ethnic group or did not specify their ethnicity and received that raw score or a higher raw score on the exam.

The second row of data within the percent section depicts the percent of candidates passing compared to the ethnic group with the highest pass rate. This information is used to determine if adverse impact exists for any of the ethnic groups. Evidence of adverse impact exists if this percentage is 80% or less.

For instance, in the example depicted above, the Pacific Islander group has the highest pass rate at 90% and therefore the pass rates of the other groups are compared with the Pacific Islander group. Five groups (White, Black, Asian, Filipino, and Native American) have adverse impact (pass rate of 80% or less) when compared with the Pacific Islander pass rate.

In order to determine the practical significance of adverse impact, the Rule of One calculation should be performed for any raw score value that is being considered for the pass point. Contact the Test Validation and Construction Unit of the State Personnel Board to receive a "how to" guide for calculating the Rule of One test.

Raw Score Beginning with the highest raw score, this column contains every raw score achieved by the candidates. It is important to review the adverse impact data in the percent section of the table for every raw score that is being considered for the pass point. It is incorrect to assume that, if adverse impact exists for one particular raw score, adverse impact will also exist for all higher raw scores.

% R Each value in this column is the percentage correct score that corresponds to the raw score values to the immediate left.

Freq This column contains the number of candidates that achieved each raw score.

Cum F This column indicates the total number of candidates that attained the particular raw score or a higher score on the exam.

In other words, this value represents, at any given score, the total number of candidates that have been accounted for. When used in setting a pass point, it provides the total number of candidates which would pass the exam if the pass point were set at that particular score.

C Frq % Each number in this column indicates the percent of candidates that achieved that raw score or a higher score on the exam. In other words, this value represents, at any given score, the percentage of candidates that have been accounted for. When used in setting a pass point, it represents the percentage of candidates which would pass the exam if the pass point were set at that particular score.

Counts by Ethnic Code This section contains the number of candidates who achieved that raw score or a higher raw score within each ethnic group.

Gender

The next adverse impact table contains adverse impact statistics for males and females. This table contains the same sections as the ethnicity adverse impact table with the exception that it is divided by gender instead of ethnicity. The procedures for interpreting this table are identical to those for interpreting the ethnicity table.

Disability

The last adverse impact table contains adverse impact statistics for those with disabilities. This table contains the same sections as the ethnicity and gender adverse impact tables with the exception that this table is divided by those with and those without disabilities. The procedures for interpreting this table are the same as those for the ethnicity and gender tables.

Frequency Histogram

The frequency histogram depicts the frequency distribution pictorially. By reviewing the histogram, you can estimate where most of the raw scores fall within the distribution.